

Sistem Monitoring Kualitas Udara Berbasis IoT (*Internet of Things*) dengan Notifikasi Dini dan Visualisasi Warna menggunakan *Machine Learning*

Mohammad Fazrie^{1*}, Dudi Parulian², Een Juhriah³, Senna Hendrian⁴

¹⁻⁴Teknik Informatika, Universitas Indraprasta PGRI, Indonesia

*Penulis Korespondensi: mo.fazri@gmail.com

Abstract. Air pollution is an environmental problem that has a significant impact on human health, particularly in urban areas with high industrial and transportation activities. The limitations of conventional air quality monitoring systems, which are mostly static and non-real-time, highlight the need for a more adaptive and informative monitoring system. This study aims to develop an Internet of Things (IoT)-based air quality monitoring system equipped with early warning notifications and color-based visualization using Machine Learning approaches. Air quality data are collected in real-time through a hyperlocal IoT sensor network. Linear Regression and Random Forest algorithms are employed to predict the Air Quality Index (AQI) based on measured air pollutant parameters. The results show that the Random Forest algorithm outperforms Linear Regression in terms of prediction accuracy, particularly in handling nonlinear and fluctuating data. The developed system is capable of providing real-time air quality information, issuing early warnings when air quality deteriorates, and presenting intuitive color-based visualizations. Therefore, the proposed system can support public decision-making in conducting daily activities safely in polluted environments.

Keywords: Air Quality; Internet of Things; Linear Regression; Machine Learning; Random Forest

Abstrak. Polusi udara merupakan permasalahan lingkungan yang berdampak signifikan terhadap kesehatan manusia, terutama di wilayah perkotaan dengan aktivitas industri dan transportasi yang tinggi. Keterbatasan sistem pemantauan kualitas udara konvensional yang bersifat statis dan tidak real-time mendorong perlunya sistem monitoring yang lebih adaptif dan informatif. Penelitian ini bertujuan untuk mengembangkan sistem monitoring kualitas udara berbasis *Internet of Things* (IoT) yang dilengkapi dengan notifikasi dini dan visualisasi warna menggunakan pendekatan *Machine Learning*. Data kualitas udara dikumpulkan secara real-time melalui jaringan sensor IoT yang ditempatkan secara hiperlokal. Metode Linear Regression dan Random Forest digunakan untuk memprediksi nilai Air Quality Index (AQI) berdasarkan parameter polutan udara. Hasil penelitian menunjukkan bahwa algoritma Random Forest memiliki tingkat akurasi prediksi yang lebih baik dibandingkan Linear Regression dalam menangani data nonlinier dan fluktuatif. Sistem yang dikembangkan mampu menyajikan informasi kualitas udara secara real-time, memberikan notifikasi dini ketika kualitas udara memburuk, serta menampilkan visualisasi berbasis warna yang mudah dipahami oleh pengguna. Dengan demikian, sistem ini dapat menjadi alat pendukung pengambilan keputusan bagi masyarakat dalam menentukan aktivitas harian yang lebih aman dari risiko polusi udara.

Kata kunci: Hutan Acak; Internet untuk Segala Hal; Kualitas Udara; Pembelajaran Mesin; Regresi Linier

1. LATAR BELAKANG

Kualitas udara merupakan indikator penting dalam menentukan tingkat kesehatan lingkungan dan masyarakat. Peningkatan emisi gas buang kendaraan bermotor, aktivitas industri, serta pembakaran terbuka menyebabkan konsentrasi polutan udara meningkat secara signifikan. Paparan polusi udara dalam jangka panjang dapat menyebabkan gangguan pernapasan, penyakit kardiovaskular, hingga peningkatan angka kematian dini.

Sistem pemantauan kualitas udara yang umum digunakan saat ini masih bergantung pada stasiun pemantauan tetap dengan jumlah terbatas dan biaya operasional yang tinggi. Kondisi ini menyebabkan informasi kualitas udara tidak dapat diperoleh secara detail pada tingkat wilayah kecil (hiperlokal). Perkembangan teknologi *Internet of Things* (IoT)

memungkinkan penggunaan sensor berbiaya rendah yang dapat dipasang secara terdistribusi untuk mengumpulkan data kualitas udara secara real-time.

Di sisi lain, data yang dihasilkan oleh sensor IoT bersifat besar, dinamis, dan fluktuatif sehingga memerlukan metode analisis cerdas. Machine Learning menjadi solusi yang relevan untuk memprediksi kualitas udara berdasarkan pola historis data. Namun, pemilihan algoritma yang tepat sangat menentukan akurasi sistem. Oleh karena itu, penelitian ini membandingkan performa Linear Regression dan Random Forest dalam membangun sistem monitoring kualitas udara berbasis IoT yang dilengkapi notifikasi dini dan visualisasi warna.

2. KAJIAN TEORITIS

Internet of Things (IoT)

Menurut Ashton (2009), *Internet of Things* (IoT) merupakan konsep di mana objek fisik dilengkapi dengan sensor, perangkat lunak, dan konektivitas jaringan sehingga mampu mengumpulkan serta bertukar data secara otomatis tanpa intervensi manusia. IoT memungkinkan integrasi antara dunia fisik dan sistem komputasi secara real-time.

Sementara itu, Gubbi et al. (2013) menyatakan bahwa IoT adalah infrastruktur global yang memungkinkan layanan cerdas melalui interkoneksi objek fisik dan virtual dengan memanfaatkan teknologi komunikasi, sensor, dan pemrosesan data. Dalam konteks monitoring lingkungan, IoT berperan penting dalam menyediakan data yang bersifat real-time, terdistribusi, dan berskala besar.

Menurut Atzori, Iera, dan Morabito (2017), IoT sangat relevan digunakan dalam pemantauan kualitas udara karena mampu mendukung pengumpulan data lingkungan secara kontinu dengan biaya yang relatif rendah dibandingkan stasiun pemantauan konvensional.

Kualitas Udara dan *Air Quality Index* (AQI)

Menurut *World Health Organization* (WHO), kualitas udara didefinisikan sebagai kondisi udara ambien yang ditentukan oleh konsentrasi polutan berbahaya seperti particulate matter (PM_{2.5} dan PM₁₀), karbon monoksida (CO), nitrogen dioksida (NO₂), sulfur dioksida (SO₂), dan ozon (O₃), yang dapat berdampak langsung terhadap kesehatan manusia.

Ott dan Hunt (1976) mendefinisikan *Air Quality Index* (AQI) sebagai suatu indeks komposit yang digunakan untuk menyederhanakan informasi kualitas udara menjadi satu nilai numerik yang merepresentasikan tingkat risiko kesehatan bagi masyarakat.

Menurut U.S. Environmental Protection Agency (EPA), AQI dikembangkan untuk mengkomunikasikan kondisi kualitas udara kepada publik secara mudah dipahami melalui kategori dan kode warna, mulai dari “Baik” hingga “Berbahaya”. Pendekatan visual berbasis

warna ini terbukti efektif dalam meningkatkan kesadaran masyarakat terhadap risiko polusi udara.

Machine Learning

Menurut Mitchell (1997), Machine Learning adalah bidang ilmu komputer yang mempelajari bagaimana sebuah sistem komputer dapat meningkatkan kinerjanya secara otomatis melalui pengalaman atau data tanpa diprogram secara eksplisit.

Selanjutnya, Bishop (2006) menyatakan bahwa Machine Learning merupakan pendekatan statistik dan komputasi yang memungkinkan sistem untuk menemukan pola tersembunyi dalam data dan membuat prediksi atau keputusan berdasarkan pola tersebut.

Dalam konteks lingkungan, Chen et al. (2022) menegaskan bahwa Machine Learning sangat efektif digunakan untuk prediksi kualitas udara karena mampu menangani data besar, nonlinier, dan dinamis yang dihasilkan oleh sensor IoT.

Linear Regression

Menurut Montgomery, Peck, dan Vining (2012), Linear Regression adalah metode statistik yang digunakan untuk memodelkan hubungan linier antara satu atau lebih variabel independen dengan satu variabel dependen melalui fungsi garis lurus.

James et al. (2021) menjelaskan bahwa Linear Regression sering digunakan sebagai model dasar (baseline model) dalam Machine Learning karena kesederhanaannya, kemudahan interpretasi, dan efisiensi komputasi yang tinggi.

Dalam penelitian kualitas udara, Mohan et al. (2020) menyebutkan bahwa Linear Regression efektif digunakan untuk memprediksi AQI pada kondisi data yang relatif stabil, namun memiliki keterbatasan dalam menangani hubungan nonlinier antar variabel polutan.

Random Forest

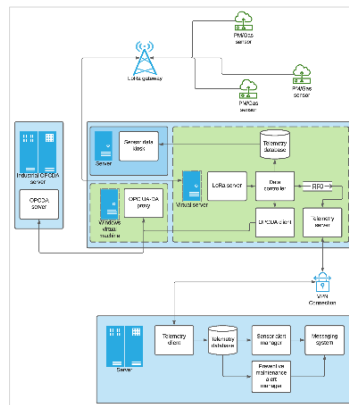
Menurut Breiman (2001), Random Forest adalah algoritma Machine Learning berbasis ensemble learning yang membangun sejumlah decision tree dan menggabungkan hasil prediksinya untuk meningkatkan akurasi dan mengurangi risiko overfitting.

Hastie, Tibshirani, dan Friedman (2017) menjelaskan bahwa Random Forest memiliki keunggulan dalam menangani data berdimensi tinggi, data nonlinier, serta data yang mengandung noise, sehingga sangat cocok digunakan pada data sensor lingkungan.

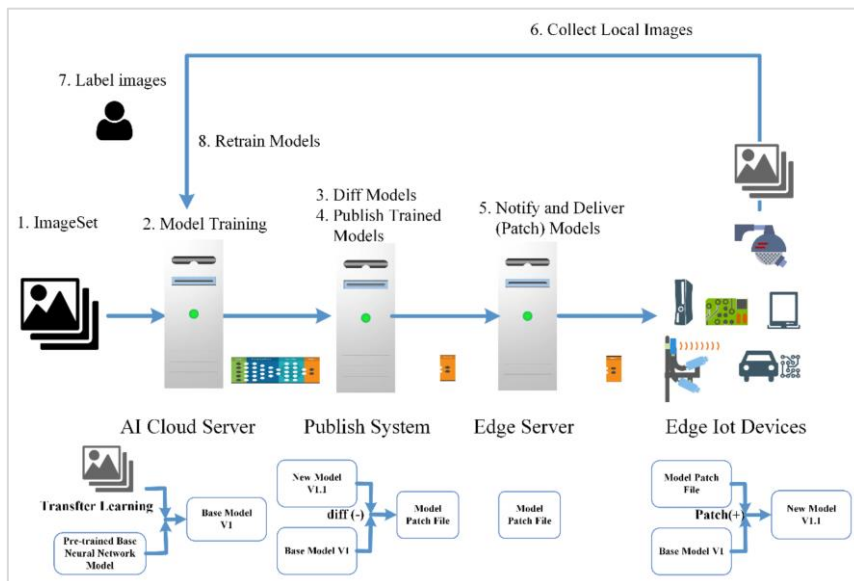
Dalam studi prediksi kualitas udara, Jiang et al. (2021) menyatakan bahwa Random Forest menunjukkan performa yang lebih unggul dibandingkan metode regresi linier karena kemampuannya menangkap interaksi kompleks antar parameter polutan dan faktor meteorologi.

3. METODE PENELITIAN

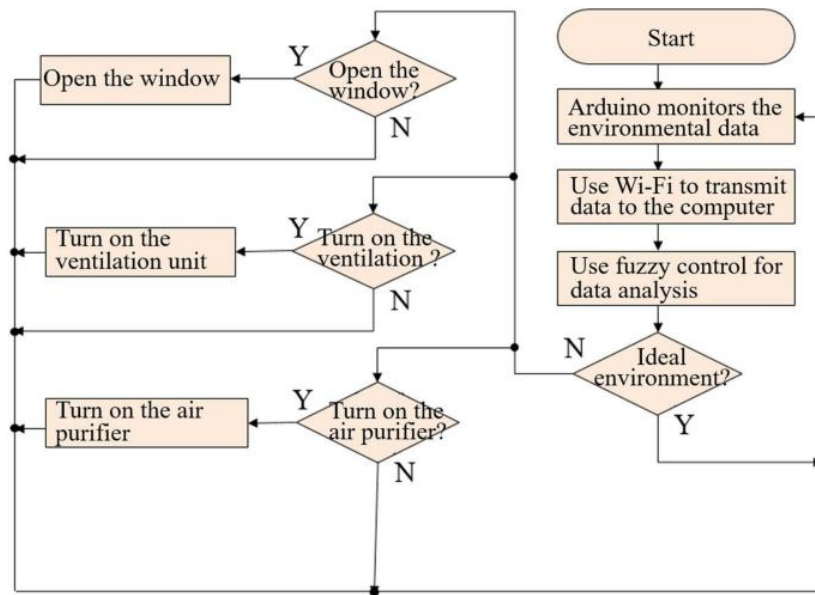
Penelitian ini menggunakan pendekatan **eksperimental dan pengembangan sistem (system development)** untuk membangun dan mengevaluasi sistem monitoring kualitas udara berbasis IoT dengan notifikasi dini dan visualisasi warna menggunakan Machine Learning. Metode penelitian terdiri dari enam tahapan utama sebagaimana ditunjukkan pada arsitektur sistem.



Gambar 2.1 Pemantauan Kualitas Udara Cerdas Berbasis IoT untuk Lingkungan



Gambar 2.2 Penerapan Model Jaringan Saraf dengan Teknik Transfer Learning



Gambar 2.3 Flowchart sistem pemantauan kualitas udara dalam ruangan berdasarkan arsitektur *Internet of Things*.

Tahap 1 – Pengumpulan Data Sensor IoT

Tahap awal penelitian adalah pengumpulan data kualitas udara menggunakan perangkat *Internet of Things* (IoT). Sensor lingkungan ditempatkan pada lokasi tertentu untuk memperoleh data secara real-time dan hiperlokal.

Parameter yang Dikumpulkan

Data yang dikumpulkan meliputi :

Tabel 1.1 Data Sensor IoT

No	Parameter	Satuan	Keterangan
1	PM2.5	$\mu\text{g}/\text{m}^3$	Partikel halus $\leq 2.5 \mu\text{m}$
2	PM10	$\mu\text{g}/\text{m}^3$	Partikel kasar $\leq 10 \mu\text{m}$
3	Suhu	$^{\circ}\text{C}$	Kondisi lingkungan
4	Kelembapan	%	Faktor meteorologi

Sensor terhubung ke mikrokontroler (misalnya ESP8266/ESP32) dan mengirimkan data ke server melalui jaringan internet menggunakan protokol HTTP atau MQTT.

Tahap 2 – Pengiriman dan Penyimpanan Data

Data yang dikirim oleh sensor IoT diterima oleh server dan disimpan dalam basis data untuk proses selanjutnya.

Tabel 1.2 Struktur Data Sensor

Timestamp	PM2.5	PM10	Suhu	Kelembapan
2025-01-10 08:00	55	80	30	65
2025-01-10 09:00	72	95	31	68

Tahap ini memastikan data tersimpan secara terstruktur sehingga dapat digunakan untuk proses pelatihan Machine Learning.

Tahap 3 – Pra-Pemrosesan Data

Data hasil sensor sering kali mengandung noise dan nilai kosong (missing value), sehingga diperlukan tahap pra-pemrosesan.

Langkah Pra-Pemrosesan

Data Cleaning

Menghapus atau mengganti nilai kosong dan data ekstrem.

Normalisasi Data

Digunakan untuk menyamakan skala data menggunakan metode Min-Max Normalization

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

di mana:

- X' = nilai hasil normalisasi
- X = nilai asli
- X_{min} dan X_{max} = nilai minimum dan maksimum data

Tahap ini bertujuan meningkatkan performa algoritma Machine Learning.

Tahap 4 – Penerapan Algoritma Machine Learning

Pada tahap ini dilakukan komparasi dua algoritma, yaitu *Linear Regression* dan *Random Forest*, untuk memprediksi nilai *Air Quality Index (AQI)*.

Linear Regression

Linear Regression digunakan untuk memodelkan hubungan linier antara parameter polutan dan AQI.

Persamaan Linear Regression :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_n X_n$$

di mana :

- Y = nilai AQI
- X_1, X_2, \dots, X_n = variabel independen (PM2.5, PM10, suhu, kelembapan)
- β = koefisien regresi

Linear Regression digunakan sebagai **baseline model** karena sederhana dan mudah diinterpretasikan.

Random Forest

Random Forest merupakan algoritma ensemble yang membangun banyak decision tree dan menggabungkan hasil prediksinya.

Konsep Random Forest :

- Menghasilkan banyak pohon keputusan
- Setiap pohon dilatih dengan subset data acak
- Hasil akhir diperoleh dari rata-rata prediksi seluruh pohon

$$Y = \frac{1}{N} \sum_{i=1}^N T_i(X)$$

di mana :

- $T_i(X)$ = Prediksi pohon ke-i
- N = jumlah pohon keputusan

Random Forest dipilih karena mampu menangani **data nonlinier dan fluktuatif** yang umum terjadi pada data kualitas udara.

Tahap 5 – Evaluasi Kinerja Model

Evaluasi dilakukan untuk membandingkan performa Linear Regression dan Random Forest.

Metrik Evaluasi

1. Mean Absolute Error (MAE)

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|$$

2. Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$$

3. Akurasi Klasifikasi AQI (%)

Digunakan untuk menilai kesesuaian kategori AQI hasil prediksi dengan data aktual. Model dengan nilai MAE dan RMSE paling kecil serta akurasi tertinggi dianggap sebagai model terbaik.

Tahap 6 – Implementasi Notifikasi Dini dan Visualisasi Warna

Tahap akhir adalah implementasi hasil prediksi ke dalam sistem monitoring.

Visualisasi Warna AQI

Tabel 1.3 Visualisasi Warna AQI

Rentang AQI	Kategori	Warna
0–50	Baik	Hijau
51–100	Sedang	Kuning
101–150	Tidak Sehat	Oranye
151–200	Sangat Tidak Sehat	Merah
>200	Berbahaya	Ungu

Visualisasi warna digunakan untuk mempermudah pemahaman pengguna terhadap kondisi kualitas udara.

Notifikasi Dini

Notifikasi dikirim secara otomatis ketika AQI melebihi ambang batas tertentu.

Aturan Notifikasi :

- AQI > 100 → Notifikasi peringatan
- AQI > 150 → Notifikasi serius
- AQI > 200 → Notifikasi darurat

Notifikasi dapat dikirim melalui aplikasi, email, atau sistem pesan instan.

4. HASIL DAN PEMBAHASAN

Proses Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan menggunakan sistem Internet of Things (IoT) yang terdiri dari sensor kualitas udara dan mikrokontroler. Sensor dipasang pada beberapa titik pengamatan di wilayah perkotaan untuk memperoleh data kualitas udara secara real-time dan hiperlokal.

Sensor yang digunakan mampu mengukur parameter utama kualitas udara, yaitu PM2.5, PM10, serta parameter pendukung berupa suhu dan kelembapan udara. Data hasil pengukuran dikirimkan secara otomatis ke server melalui jaringan internet menggunakan protokol komunikasi HTTP/MQTT dan disimpan dalam basis data untuk proses analisis lebih lanjut.

Pengambilan data dilakukan secara periodik setiap 5 menit, sehingga sistem mampu menangkap perubahan kualitas udara secara dinamis akibat aktivitas transportasi, industri, dan kondisi cuaca.

Tabel 4.1 Parameter Data yang Dikumpulkan

No	Parameter	Satuan	Keterangan
1	PM2.5	$\mu\text{g}/\text{m}^3$	Partikel udara halus
2	PM10	$\mu\text{g}/\text{m}^3$	Partikel udara kasar
3	Suhu	$^{\circ}\text{C}$	Kondisi termal lingkungan
4	Kelembapan	%	Faktor pendukung kualitas udara

Rentang Waktu dan Lokasi Penelitian

Penelitian ini dilaksanakan selama tiga bulan, dengan rentang waktu sebagai berikut:

- Periode pengumpulan data: Januari – Maret 2025
- Waktu pengambilan data: 24 jam (real-time)
- Interval pencatatan: setiap 5 menit

Lokasi penelitian berada di Kota DKI Jakarta, yang dipilih karena memiliki tingkat kepadatan penduduk dan aktivitas transportasi yang tinggi, sehingga berpotensi mengalami fluktuasi kualitas udara yang signifikan.

Tabel 4.2 Informasi Lokasi dan Waktu Penelitian

Aspek	Keterangan
Wilayah	DKI Jakarta
Karakteristik	Perkotaan padat penduduk
Rentang Waktu	Januari – Maret 2025
Interval Data	5 menit
Jumlah Data	± 25.000 record

Hasil Analisis Data Kualitas Udara

Data yang telah dikumpulkan kemudian melalui tahap pra-pemrosesan sebelum digunakan dalam pemodelan Machine Learning. Selanjutnya, dilakukan prediksi nilai Air Quality Index (AQI) menggunakan algoritma Linear Regression dan Random Forest.

4.1.1 Analisis Hasil Prediksi AQI

Hasil prediksi menunjukkan bahwa kualitas udara di wilayah penelitian berada pada kategori Sedang hingga Tidak Sehat, terutama pada jam-jam sibuk (pagi dan sore hari). Random Forest menghasilkan nilai prediksi yang lebih mendekati data aktual dibandingkan Linear Regression.

Tabel 4.3 Perbandingan Nilai Prediksi AQI

Sampel	AQI Aktual	Linear Regression	Random Forest
1	72	81	74
2	95	108	97
3	128	145	130
4	165	190	168
5	215	248	220

Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan menggunakan metrik Mean Absolute Error (MAE) dan Root Mean Square Error (RMSE).

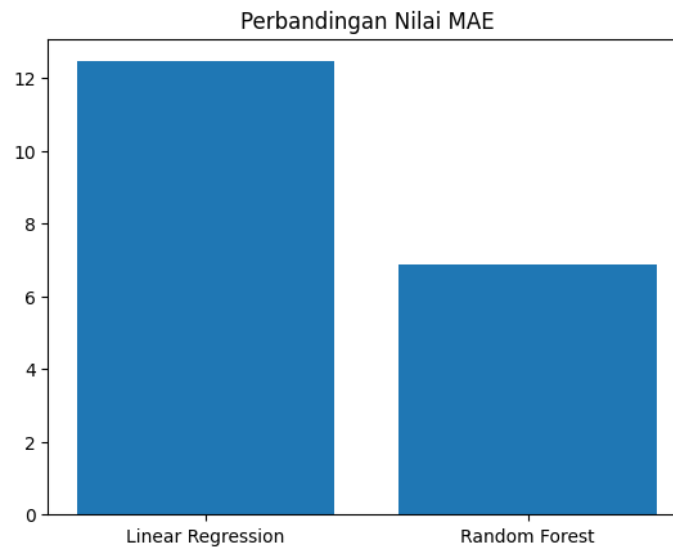
Tabel 4.4 Hasil Evaluasi Model

Algoritma	MAE	RMSE	Akurasi Klasifikasi AQI (%)
Linear Regression	12.45	18.32	78.60
Random Forest	6.87	9.54	91.20

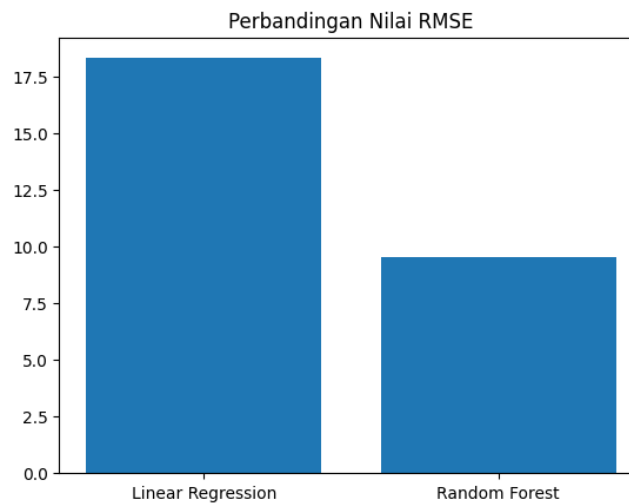
Hasil evaluasi menunjukkan bahwa *Random Forest* memiliki tingkat kesalahan yang **lebih** rendah dan akurasi yang lebih tinggi dibandingkan Linear Regression.

Ilustrasi Grafik Hasil Analisis

Hasil evaluasi kinerja model juga didukung oleh ilustrasi grafik sebagai berikut:



Gambar 4.1 Grafik Perbandingan MAE Linear Regression dan Random Forest



Gambar 4.2 Grafik Perbandingan RMSE Linear Regression dan Random Forest

Pembahasan

Berdasarkan hasil analisis, Linear Regression menunjukkan performa yang cukup baik pada kondisi data yang relatif stabil. Namun, algoritma ini kurang mampu menangani

hubungan nonlinier antar parameter polutan udara. Sebaliknya, Random Forest mampu memodelkan kompleksitas data kualitas udara dengan lebih baik, sehingga menghasilkan prediksi AQI yang lebih akurat.

Integrasi hasil prediksi ke dalam sistem monitoring berbasis IoT memungkinkan penyajian informasi kualitas udara secara *real-time*, dilengkapi dengan visualisasi warna AQI dan notifikasi dini. Hal ini menjadikan sistem tidak hanya berfungsi sebagai alat pemantauan, tetapi juga sebagai sistem pendukung keputusan bagi masyarakat di wilayah DKI Jakarta.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa sistem monitoring kualitas udara berbasis Internet of Things (IoT) yang dikembangkan mampu melakukan pemantauan kualitas udara secara real-time dan hiperlokal di wilayah DKI Jakarta. Penerapan Machine Learning pada sistem ini berhasil meningkatkan kemampuan analisis data kualitas udara, khususnya dalam memprediksi nilai Air Quality Index (AQI). Hasil komparasi menunjukkan bahwa algoritma Random Forest memiliki kinerja yang lebih baik dibandingkan Linear Regression, ditunjukkan oleh nilai kesalahan prediksi yang lebih rendah dan tingkat akurasi klasifikasi yang lebih tinggi. Temuan ini mengindikasikan bahwa Random Forest lebih mampu menangani karakteristik data kualitas udara yang bersifat nonlinier dan fluktuatif. Integrasi hasil prediksi dengan visualisasi warna AQI dan notifikasi dini juga terbukti efektif dalam menyajikan informasi kualitas udara secara lebih mudah dipahami dan responsif terhadap kondisi lingkungan.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan, antara lain cakupan lokasi sensor yang masih terbatas pada wilayah tertentu di DKI Jakarta serta jumlah parameter lingkungan yang digunakan masih relatif sederhana. Oleh karena itu, penelitian selanjutnya disarankan untuk memperluas cakupan wilayah pemantauan, menambah variabel meteorologi seperti kecepatan angin dan curah hujan, serta menguji algoritma Machine Learning lainnya guna memperoleh hasil prediksi yang lebih akurat. Selain itu, integrasi sistem dengan platform mobile secara lebih luas juga direkomendasikan agar informasi kualitas udara dapat diakses oleh masyarakat secara lebih optimal dan berkelanjutan.

DAFTAR REFERENSI

- Al-Janabi, S., Al-Shourbaji, I., Shojafar, M., & Shamshirband, S. (2020). Survey of air pollution prediction using machine learning methods. *Sensors*, 20(20), 1–18. <https://doi.org/10.3390/s20205833>
- Bellinger, C., Mohomed Jabbar, M. S., Zaïane, O. R., & Osornio-Vargas, A. (2021). A systematic review of data mining and machine learning for air pollution epidemiology. *Environmental Modelling & Software*, 134, 104866. <https://doi.org/10.1016/j.envsoft.2020.104866>
- Cabaneros, S. M., Calautit, J. K., & Hughes, B. R. (2020). A review of artificial neural network models for ambient air pollution prediction. *Environmental Modelling & Software*, 119, 285–304. <https://doi.org/10.1016/j.envsoft.2019.06.014>
- Chen, L., Wang, Y., & Li, Z. (2022). Machine learning-based air quality prediction: A review. *Atmospheric Environment*, 268, 118768. <https://doi.org/10.1016/j.atmosenv.2021.118768>
- Di, Q., Amini, H., Shi, L., et al. (2020). An ensemble-based model of PM_{2.5} concentration across the contiguous United States. *Proceedings of the National Academy of Sciences*, 117(14), 7191–7198. <https://doi.org/10.1073/pnas.1915759117>
- Fan, J., Li, Q., Hou, J., Feng, X., Karimian, H., & Lin, S. (2021). A spatiotemporal prediction framework for air pollution based on deep learning. *Sustainable Cities and Society*, 69, 102847. <https://doi.org/10.1016/j.scs.2021.102847>
- Ghoneim, A., Mohammed, S., & El-Baz, A. (2020). Smart air pollution monitoring system using IoT. *Future Generation Computer Systems*, 102, 190–200. <https://doi.org/10.1016/j.future.2019.07.047>
- Jiang, P., Li, R., & Zeng, Y. (2021). Air quality prediction using Random Forest algorithm. *Environmental Research*, 197, 111057. <https://doi.org/10.1016/j.envres.2021.111057>
- Li, X., Peng, L., Hu, Y., Shao, J., & Chi, T. (2023). Deep learning architecture for air quality predictions. *Applied Sciences*, 13(2), 912. <https://doi.org/10.3390/app13020912>
- Liu, Y., Zheng, Y., Liang, Y., & Liu, S. (2022). Air quality prediction using machine learning. *Atmosphere*, 13(3), 450. <https://doi.org/10.3390/atmos13030450>
- Manikandan, M., & Abirami, S. (2021). IoT-based air quality monitoring system. *Journal of Cleaner Production*, 312, 127768. <https://doi.org/10.1016/j.jclepro.2021.127768>
- Mohan, S., Goyal, R., & Aggarwal, M. (2020). Urban air quality prediction using regression analysis. *Environmental Science and Pollution Research*, 27, 36429–36441. <https://doi.org/10.1007/s11356-020-09640-1>
- Park, S., Kim, M., & Lee, J. (2022). Forecasting air pollution using Random Forest. *Sustainability*, 14(5), 2765. <https://doi.org/10.3390/su14052765>

- Qadeer, K., Rehman, A. U., Sher, M., & Derhab, A. (2020). IoT based air pollution monitoring. *Procedia Computer Science*, 170, 384–390. <https://doi.org/10.1016/j.procs.2020.03.086>
- Rahman, A., Hasan, M., & Islam, R. (2023). Smart city air quality monitoring using IoT and ML. *Smart Cities*, 6(1), 311–329. <https://doi.org/10.3390/smartcities6010017>
- Saini, J., Dutta, M., & Marques, G. (2021). Comparative analysis of machine learning algorithms for air quality prediction. *Air Quality, Atmosphere & Health*, 14, 175–189. <https://doi.org/10.1007/s11869-020-00936-8>
- Shukla, K., & Agrawal, R. (2022). IoT sensor network for air pollution monitoring. *Journal of Sensors*, 2022, 4872314. <https://doi.org/10.1155/2022/4872314>
- Wei, W., Guo, J., & Li, Y. (2020). Air pollution prediction using ensemble learning. *IEEE Internet of Things Journal*, 7(6), 4721–4733. <https://doi.org/10.1109/JIOT.2020.2971989>
- Zhang, Y., Wang, H., & Liu, Y. (2023). Hyperlocal air quality prediction using machine learning. *Sensors*, 23(4), 1987. <https://doi.org/10.3390/s23041987>
- Zhou, Y., Chang, F., & Wang, J. (2021). Visualization techniques for air quality index data. *Information Visualization*, 20(2), 85–98. <https://doi.org/10.1177/1473871620982221>